**AFRL-RI-RS-TR-2009-73**
**Final Technical Report**
**March 2009**

# WIRELESS TECHNOLOGY

The State University of New York at Buffalo

***APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED***.

**STINFO COPY**

**AIR FORCE RESEARCH LABORATORY**
**INFORMATION DIRECTORATE**
**ROME RESEARCH SITE**
**ROME, NEW YORK**

# NOTICE AND SIGNATURE PAGE



AFRL-RI-RS-TR-2009-73 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

FOR THE DIRECTOR:

| | |
|---|---|
| /s/ | /s/ |
| STEPHEN REICHHART<br>Work Unit Manager | WARREN H. DEBANY, JR.<br>Technical Advisor, Information Grid Division<br>Information Directorate |

# REPORT DOCUMENTATION PAGE

*Form Approved*
**OMB No. 0704-0188**



**1. REPORT DATE** *(DD-MM-YYYY)*
MAR 09

**2. REPORT TYPE**
Final

**3. DATES COVERED** *(From - To)*
May 03 – Sep 08

**4. TITLE AND SUBTITLE**
WIRELESS TECHNOLOGY

**5a. CONTRACT NUMBER**

**5b. GRANT NUMBER**
F30602-03-2-0105

**5c. PROGRAM ELEMENT NUMBER**
62702F

**5d. PROJECT NUMBER**
WIRE

**5e. TASK NUMBER**
TE

**5f. WORK UNIT NUMBER**
CH

**6. AUTHOR(S)**
Stella N. Batalama

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**
The State University of New York at Buffalo
520 Lee Entrance
Amherst NY 14228

**8. PERFORMING ORGANIZATION REPORT NUMBER**

**9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)**
AFRL/RIGC
525 Brooks Rd.
Rome NY 13441-4505

**10. SPONSOR/MONITOR'S ACRONYM(S)**

**11. SPONSORING/MONITORING AGENCY REPORT NUMBER**
AFRL-RI-RS-TR-2009-73

**12. DISTRIBUTION AVAILABILITY STATEMENT**
*APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED. PA# 88ABW-2009-0986*

**13. SUPPLEMENTARY NOTES**

**14. ABSTRACT**
This report covers research on the subject of generalized likelihood radio test packet data detectors. The research led to the design of a novel generalized likelihood ratio test (GLRT)-type packet-data detectors for general multi-access/multi-user digital communication systems and analytical performance evaluation tools for finite data packet sizes. Developed for the known channel case was a coherent GLRT packet-data detector, while for the unknown channel case two detectors were developed. For the unknown channel case a coherent pilot assisted GLRT packet-data detector and a differential phase-shift-keying (DPSK) GLRT packet-data detector. Efficient suboptimum implementations of the above schemes that exhibit complexity linear in the packet size were also considered. Simulation studied evaluated the performance of the proposed schemes in the context of packet-data code-division multiple access (CDMA) communications.

**15. SUBJECT TERMS**
Coherent packet-data detection, multi-access/multi-user communication

| 16. SECURITY CLASSIFICATION OF: a. REPORT | b. ABSTRACT | c. THIS PAGE | 17. LIMITATION OF ABSTRACT | 18. NUMBER OF PAGES | 19a. NAME OF RESPONSIBLE PERSON |
|---|---|---|---|---|---|
| U | U | U | UU | 42 | Stephen Reichhart |
| | | | | | 19b. TELEPHONE NUMBER (*Include area code*) N/A |

**Standard Form 298 (Rev. 8-98)**
Prescribed by ANSI Std. Z39.18

# Table of Contents

# List of Figures

# 1. Summary

My first administrative task as the Acting Director of CITE was to coordinate the research activities of CITE's in-house researchers and external collaborators and prepare group presentations and demonstrations during the Scientific Advisory Board (SAB) review of the Information Directorate (IF) that was held in November 2003. The SAB review is a biannual 2-day event where a team of high-profile external evaluators (from government, academia, and industrial sector) is visiting on-site to review and evaluate the performance of the Directorate. As part of IF, CITE participated in the SAB review with an overarching presentation of its research activities and accomplishments that included 11 research posters/presentations and 3 experimental demonstrations. In the SAB final report, the evaluation of CITE includes the following statement: "IF, and AFRL more generally, should consider CITE as a candidate best practice model to focus research by leveraging academia, industry, and internal activities."

In this context, during my 11-month tenure as the Acting Director of CITE I addressed and acted upon four items: CITE research focus, personnel needs, budget/funding, in-house and external collaborators.

*(i)Research Focus*

The research activities of CITE initially evolved around video-centric R&D. In terms of future growth, at that time, CITE was moving toward broadening its horizon and R&D scope from video-centric to (a) general (covert, secure, or open) transmission and (b) detection and exploitation of all forms of digital signals (audio, image/video, data) that is military specific. The administrative action that I took regarding this item was to organize a strategic planning

meeting that involved seven external high profile evaluators and seven AFRL participants (administrators and research scientists including the Chief Scientist of IF). The Strategic Planning Committee concluded that at a high level of abstraction there are four primary theoretical research areas of interest to CITE; namely, signal design, scalability, adaptivity, and nonstationarity. The signal design area focuses on avoiding interference by selective modulation, code design for signal multiplexing, and adaptive coding strategies. Scalability is viewed as the means to increase average information content of the transmitted signal by matching the bandwidths of transmitter, channel, and receiver. Adaptivity is considered in the context of minimizing quickly and effectively the effects of interference, natural or man-made, on the transmission, collection, and exploitation of information. Finally, within the nonstationarity focus area, the objective is to characterize and exploit the time-varying nature of transmission and collection environments (e.g., small sample support adaptivity can exploit local stationarity). The developments in these theoretical research areas are applied to transmission exploitation applications. Examples of projected needs of the next generation warfighter in the areas of transmission and exploitation include:

– Wireless digital connectivity beyond LoS; "communicate with any asset any time."

– Locate, intercept, and exploit digital information quickly and reliably.

– Automate exploitation process to assist the human operator/analyst (targeting, damage assessment, annotation, summarization).

– Develop miniature wireless sensor packages with daylight imaging, ranging designation, and GPS.

– Develop SW/HW radios that combine multiple data link waveforms and protocols.

The Strategic Planning committee also concluded that the T & E critical technical areas that

CITE should focus on are: All-IP wireless (basic research component), video over wireless IP (applied research component), and sensor networks (target application/testbed). The above areas encompass research on next generation multiple access technologies, code/signature design and adaptive assignment for code division multiplexing, data authentication/information assurance, interference mitigation/multiuser detection for effective data collection and reliable data exploitation, surveillance, eavesdropping, covert communications, and cross-layer end-to-end optimization.

In terms of maintaining an evaluation/demonstration platform of emerging technologies, the Strategic Planning Committee suggested that CITE should maintain (and continuously upgrade) a unifying evaluation/demonstration vehicle (testbed) in order to facilitate technology forecasting, feasibility analysis and technology transition, provide cost vs. return preliminary assessment, foster cooperative research among academia, industry, and government, provide a unique edge in attracting collaborators, support Directorate's technical vision, and serve as a multiservice evaluation center for military information transmission and exploitation.

In agreement with or in response to the Strategic Planning Meeting findings I took the following actions:

− Initiated two joint projects with IFEC on wideband multiuser detection for effective data exploitation in high density rapidly changing environments (funded by Tactical Sigint Technology Program).

− Initiated a dialog (through a series of lectures and meetings) with IFEC for transition of CITE's theoretical developments to audio-specific problems (accent identification, speech/language recognition).

− Initiated/funded a joint industry-academia-in-house project for the development of a unifying

layered video transmission evaluation platform.

*(ii) Personnel, Budget/funding, Collaborators*

Early on in my tenure as the Acting Director of CITE, I realized that there was a need for new hirings with competitive salaries and decreased management responsibilities to build a technical research core that fosters collaboration among T & E in-house teams, and academic and industrial research groups. My administrative actions with respect to this item was to seek and obtain authorization to recruit, initiate the advertisement of 6.1 research job openings, and, finally, present the top applicants to the IF administration. The IF administration then proceeded with job offers with very competitive salary and limited program management responsibilities.

The administrative action that I took was the initiation of a joint project with IFEC on wideband multiuser detection for effective data exploitation in high density rapidly changing environments.

Finally, the administrative actions that I took to serve better CITE's collaborators and sponsors was (*i*) to propose that CITE starts funding selected AFOSR supported University research (that is within the scope of CITE) to transition research to CITE's testbed and (*ii*) to initiate/fund a joint industry-academia in-house project for the development of a unifying evaluation/demonstration platform of layered video transmission over multirate SS wireless channels with embedded interference resistant receiver technology.

Apart from the administrative responsibilities as the Acting Director of CITE, during my sabbatical leave and the years that followed I contacted research work on the subject of generalized likelihood ratio test packet data detectors. Initial findings were summarized and published to the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)

that was held in Montreal in May 2004. The paper was accepted for publication in the conference proceedings and was presented at a technical session of the conference. Subsequently, the full body of this work of this work was published in the IEEE Transactions on Communications, in the Feb. issue of 2008 (Volume 56, Page(s):222 - 233). In this work, we designed novel generalized likelihood ratio test (GLRT)-type packet-data detectors for general multiaccess/multiuser digital communication systems and we developed analytical performance evaluation tools for finite data packet sizes. For the known channel case we derived a coherent GLRT packet-data detector, while for the unknown channel case we derived both a coherent pilot assisted GLRT packet-data detector and a differential phase-shift-keying (DPSK) GLRT packet-data detector. Efficient suboptimum implementations of the above schemes that exhibit complexity linear in the packet size were also considered. Simulation studies evaluated the performance of the proposed schemes in the context of packet-data code-division multiple access (CDMA) communications. The rest of this report contains a description of the above research developments in some detail. Further details and additional research activities have been reported in several progress reports already submitted to AFRL during the period Aug. 2004-Sept. 2008.

# 2. Introduction

The optimum rule for the detection of a transmitted packet of digital data under perfectly known parameters of the received data joint probability density function is the well known likelihood ratio test (LRT) that selects the most likely data combination among the finite set of alternatives. When, however, there are unknown parameters in the received signal model/density and a uniformly most powerful (UMP) test does not exist [1], the design of a detection scheme

becomes a coupled optimization process that involves joint detection (hypothesis testing) and parameter estimation. As such, we can either solve the estimation part first (i.e., maximize the likelihood of each hypothesis with respect to the unknown parameters) and then solve the detection part (i.e., choose the most likely hypothesis) or execute the optimization sequence in the opposite order. Under certain general conditions, the above double maximization problems are equivalent and result to what is known as the generalized likelihood ratio test (GLRT). In particular, the estimation-detection sequence of optimization is the most intuitive and results in a likelihood ratio test that utilizes maximum likelihood (ML) estimates of the unknown parameters. On the other hand, for specific applications the detection-estimation optimization sequence, although a more difficult optimization problem in general, may lead to computationally simpler test implementations than the estimation-detection sequence [2]. In any case, the overall statistical optimality of GLRT tests is difficult to be determined theoretically, if at all possible. An alternative, ad-hoc but frequently used approach to the design of a detection scheme in the presence of unknown parameters in the distribution of the received data is to proceed by directly substituting parameter estimates in the LRT formula.

GLRT has been a rather popular methodology in the past for radar signal detection problems (as seen for example in [3]– [6] and references therein) while it has been given limited consideration in the context of multiaccess/multiuser digital communications [7], [8]. The binary nature of the radar hypothesis testing problem as well as the availability of secondary data (pure disturbance observations) in addition to primary data (data that include both the signal of interest and disturbance) facilitates greatly the design of the GLRT test. On the other hand, we may argue that GLRT approaches to multiaccess/multiuser digital communications are not so straightforward because of the usual non-binary nature of the detection problem, the

absence of secondary data, and/or the non-Gaussian characteristics of the disturbance (e.g., multiaccess interference).

In this work we propose novel GLRT packet-data detectors for general multiaccess/multiuser digital communication systems. In particular, we develop: *a)* a coherent GLRT packet-data detector for the known channel case, *b)* a coherent GLRT pilot assisted detector for the unknown channel case (the channel is estimated implicitly as part of the GLRT formulation while short pilot signaling is used to resolve phase ambiguity), and *c)* a differential GLRT detector for differentially encoded packet-data. In view of the exponential complexity in the size of the data packet of the above GLRT schemes, we also propose suboptimum implementations that exhibit linear complexity. Last but not least, we develop analytical performance evaluation tools for finite data packet sizes. The importance/novelty of our analytical performance evaluation tools lies in the fact that they deviate from the conventional and convenient yet inaccurate performance analysis assumption of infinitely long packet sizes. Instead, our formulas provide the bit error rate (BER) that can be reached by a GLRT test for a given finite data packet size as well as the size of a data packet that is necessary for the test to reach a pre-specified BER level. Comparative studies included in this work establish analytically the BER performance merits of the new GLRT tests relative to the popular practice of directly substituting sample average estimates of the unknown parameters in the LRT formula. The proposed GLRT detectors are evaluated in the context of packet-data CDMA communications and state-of-the-art performance is demonstrated.

The rest of the report is organized as follows. Our signal model and some background information are presented in Section III. The known and unknown channel cases are treated in Section IV-A and IV-C, respectively. Therein, the proposed packet-data detectors are derived,

their suboptimum implementations of linear complexity are outlined, and the relationship between the size of the available data record and the BER performance that can be reached by each GLRT scheme is identified analytically. Section IV-C is devoted to simulation studies and comparisons. A few concluding remarks are drawn in Section V.

# 3. Methods, Assumptions, and Procedures

Consider the following general discrete-time signal model for a transmitted data packet of interest of size $N$:

$$\mathbf{x}_i = \sqrt{E} b_i \mathbf{s}, \quad i = 1, 2, \ldots, N, \tag{1}$$

where the information bits $b_i$, $i = 1, \ldots, N$, take values $\pm 1$ with equal probability, are independent from each other, and modulate a known $G$-dimension discrete-time complex signal waveform of unit norm, $\mathbf{s} \in \mathbb{C}^G$, $||\mathbf{s}|| = 1$. With this setup, $E$ represents total transmitted energy per bit. We assume that the transmitted signal experiences multipath quasi-static fading of maximum delay $M < G$ sampling periods with negligible inter-symbol interference (ISI) effects. If $\{\mathbf{y}_i \in \mathbb{C}^L\}_{i=1}^N$, $L = G + M$, denotes the corresponding received data packet, then

$$\mathbf{y}_i = b_i \sqrt{E} \mathbf{S} \mathbf{a} + \mathbf{n}_i, \quad i = 1, \ldots, N, \tag{2}$$

where $\mathbf{S}$ is the $L \times M$ known signal waveform delay matrix, $\mathbf{a} \in \mathbb{C}^M$ is the vector of the multipath channel coefficients that are assumed to remain constant during the transmission of the data packet, and $\mathbf{n}_i$, $i = 1, \ldots, N$, is a sequence of independent identically distributed zero mean Gaussian vector with *unknown* covariance matrix $\mathbf{R}_n$, that represent comprehensively channel interference and noise that is independent of the data sequence $b_i$, $i = 1, \ldots, N$.

The probability density function (pdf) of the observations $\mathbf{Y} \stackrel{\triangle}{=} [\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_N]$ conditioned on the transmitted bits $\mathbf{b} = [b_1, b_2, \ldots, b_N]^T$ can be expressed in the following compact form:

$$f\left(\mathbf{Y} \,|\mathbf{b}; E, \mathbf{S}, \mathbf{a}, \mathbf{R}_n\right) = \frac{1}{\pi^{LN} \left|\mathbf{R}_n\right|^N} e^{\mathrm{trace}\left[-\mathbf{R}_n^{-1}\left(\mathbf{Y}-\sqrt{E}\mathbf{Sab}^T\right)\left(\mathbf{Y}-\sqrt{E}\mathbf{Sab}^T\right)^H\right]}. \tag{3}$$

When $\mathbf{a}$ and $\mathbf{R}_n$ are perfectly known, the optimum rule for the detection of $b_i$, $i = 1, \ldots N$, simplifies to the familiar one-shot tests

$$\widehat{b}_{i_{ML}} = \mathrm{sgn}\left[\mathrm{Re}\left\{\left(\mathbf{R}_n^{-1}\mathbf{Sa}\right)^H \mathbf{y}_i\right\}\right], \ \ i = 1, \ldots N, \tag{4}$$

where sgn$[x]$ extracts the sign of the real variable $x$ and Re$(y)$ extracts the real part of the complex scalar $y$. In other words, the optimum maximum likelihood (ML) detector of $\mathbf{b} \in \{\pm 1\}^N$ reduces to N individual applications of the linear filter $\mathbf{R}_n^{-1}\mathbf{Sa}$ followed by an output sign detector and the detection of all bits in the packet has computational cost linear in the packet size (detection of $b_i$ depends only on $\mathbf{y}_i$ and joint detection of all bits in the packet is equivalent to disjoint bit-by-bit detection).

On the other hand, when a priori knowledge of the parameters $\mathbf{a}$ and/or $\mathbf{R}_n$ cannot be assumed, we may proceed in one of two different ways: *(i)* We can use again the parametrically described test in (4) and substitute the unknown parameters/statistics by corresponding estimates (usually sample-average), which results in a scheme that maintains linear complexity in the packet size; or *(ii)* we may carry out joint detection and parameter estimation which results in superior performance GLRT schemes at the expense of increased complexity. In the next section, we propose a novel packet-data GLRT test for the case of known multipath channels (parameter $\mathbf{a}$) and colored Gaussian disturbance of unknown correlation matrix (parameter $\mathbf{R}_n$) and we establish analytically how the new GLRT test compares to the common sample-average LRT parameter substitution approach. Alongside, we develop analytical tools that provide the

BER performance of the proposed test for any given data packet size as well as the data packet size that is necessary for the test to reach a given BER level.

# 4. Results and Discussion

## 4.1 GLRT Detection: Known Channel

### *4.1.1 Algorithmic Development*

For convenience, define $\mathbf{v} \stackrel{\triangle}{=} \sqrt{E}\mathbf{S}\mathbf{a}$ where $\mathbf{a}$ is the known channel coefficient vector and $\mathbf{S}$ is the known signal waveform matrix. The GLRT packet-data detector is given by the following Proposition.

**Proposition 1** *The GLRT test for the detection of the data packet* $\mathbf{b}$ *of size* $N$ *in the presence of complex Gaussian disturbance with unknown covariance matrix* $\mathbf{R}_n$ *is*

$$\widehat{\mathbf{b}}_{GLRT} = \arg\max_{\mathbf{b}} \left\{ \max_{\mathbf{R}_n} f\left(\mathbf{Y} \left| \mathbf{b}, \mathbf{v}, \mathbf{R}_n \right.\right) \right\} = \arg\max_{\mathbf{b}} \, l_1(\mathbf{b}) \tag{5}$$

*where*

$$\begin{aligned} l_1(\mathbf{b}) \stackrel{\triangle}{=} \; & N\mathbf{b}^T\mathbf{Y}^H \left[\mathbf{R}_{SA}(N)\right]^{-1}\mathbf{v} + N\mathbf{v}^H\left[\mathbf{R}_{SA}(N)\right]^{-1}\mathbf{Y}\mathbf{b} + \left(\mathbf{b}^T\mathbf{Y}^H\left[\mathbf{R}_{SA}(N)\right]^{-1}\mathbf{Y}\mathbf{b}\right)\left(\mathbf{v}^H\left[\mathbf{R}_{SA}(N)\right]^{-1}\mathbf{v}\right) \\ & - \left(\mathbf{b}^T\mathbf{Y}^H\left[\mathbf{R}_{SA}(N)\right]^{-1}\mathbf{v}\right)\left(\mathbf{v}^H\left[\mathbf{R}_{SA}(N)\right]^{-1}\mathbf{Y}\mathbf{b}\right) \end{aligned} \tag{6}$$

*and* $\mathbf{R}_{SA}(N) \stackrel{\triangle}{=} \frac{1}{N}\mathbf{Y}\mathbf{Y}^H$ *is the sample average received data correlation matrix.*

*Proof:* For a given bit combination $\mathbf{b}$, the maximum of $f\left(\mathbf{Y} \left| \mathbf{b}, \mathbf{v}, \mathbf{R}_n \right.\right)$ is reached when $\mathbf{R}_n$ is the ML estimate, i.e. $\mathbf{R}_n = \widehat{\mathbf{R}}_{n_{ML}}(\mathbf{b}) = \frac{1}{N}\left(\mathbf{Y} - \mathbf{v}\mathbf{b}^T\right)\left(\mathbf{Y} - \mathbf{v}\mathbf{b}^T\right)^H$. Thus, (5) reduces readily to

$$\widehat{\mathbf{b}}_{GLRT} = \arg\min_{\mathbf{b}} \left| \frac{1}{N}\left(\mathbf{Y} - \mathbf{v}\mathbf{b}^T\right)\left(\mathbf{Y} - \mathbf{v}\mathbf{b}^T\right)^H \right| = \arg\min_{\mathbf{b}} \left| \mathbf{R}_{SA}(N) - \frac{\mathbf{v}\mathbf{b}^T\mathbf{Y}^H}{N} - \frac{\mathbf{Y}\mathbf{b}\mathbf{v}^H}{N} + \mathbf{v}\mathbf{v}^H \right| . \tag{7}$$

The test in (7) can be further reduced to (5) using the identity

$$\left|\mathbf{R}_{SA}-\mathbf{v}\left(\frac{\mathbf{Yb}}{N}\right)^{H}-\frac{\mathbf{Yb}}{N}\mathbf{v}^{H}+\mathbf{v}\mathbf{v}^{H}\right| = |\mathbf{R}_{SA}|\left[1+\mathbf{v}^{H}\mathbf{R}_{SA}^{-1}\mathbf{v}-\mathbf{v}^{H}\mathbf{R}_{SA}^{-1}\frac{\mathbf{Yb}}{N}-\left(\frac{\mathbf{Yb}}{N}\right)^{H}\mathbf{R}_{SA}^{-1}\mathbf{v}\right.$$
$$\left.+\mathbf{v}^{H}\mathbf{R}_{SA}^{-1}\frac{\mathbf{Yb}}{N}\left(\frac{\mathbf{Yb}}{N}\right)^{H}\mathbf{R}_{SA}^{-1}\mathbf{v}-\mathbf{v}^{H}\mathbf{R}_{SA}^{-1}\mathbf{v}\left(\frac{\mathbf{Yb}}{N}\right)^{H}\mathbf{R}_{SA}^{-1}\frac{\mathbf{Yb}}{N}\right] \quad (8)$$

that holds true for any Hermitian positive definite matrix $\mathbf{R}_{SA}$ and vectors $\mathbf{v}$ and $\frac{\mathbf{Yb}}{N}$. ■

We note that direct implementation of test in (5) has complexity exponential in the packet size $N$.

The GLRT test in (5) can be contrasted with the standard "sample-matrix-inversion" (SMI) detection scheme, that replaces the unknown parameter $\mathbf{R}_n$ of the LRT expression in (4) by a sample average estimate $\mathbf{R}_{n_{SA}}(K) \triangleq \frac{1}{K}\sum_{k=1}^{K}\mathbf{n}_k\mathbf{n}_k^H$ based on pure disturbance observations $\mathbf{n}_k$, $k=1,\ldots,K$ that are independent from $\mathbf{y}_i$, $i=1,\ldots,N$:

$$\widehat{b}_{i_{SMI}} = \text{sgn}\left[\text{Re}\left\{\mathbf{v}^{H}\mathbf{R}_{n_{SA}}^{-1}(K)\mathbf{y}_i\right\}\right], \quad i=1,\ldots,N. \quad (9)$$

When pure disturbance observations (secondary data) $\mathbf{n}_k$, $k=1,\ldots,K$ are not available, a popular version of the test in (9) utilizes directly the sample-average correlation matrix of the (desired-signal-present) received data, $\mathbf{R}_{SA}(N)$, evaluated using the *same* received data $\mathbf{y}_i$, $i=1,\ldots,N$, that are processed by the detector. We denote this test by

$$\widehat{b}_{i_{SMI-dsp}} = \text{sgn}\left[\text{Re}\left\{\mathbf{v}^{H}\mathbf{R}_{SA}^{-1}(N)\mathbf{y}_i\right\}\right], \quad i=1,\ldots,N, \quad (10)$$

where the subscript part "dsp" stands for desired-signal-present. While it is understood that (9) and (10) converge with probability 1 to the test in (4) as the $K,N\to\infty$, recent analytical results on short data record adaptive filtering [9]– [12] indicate that for finite sample support of equal size ($K=N$) the test in (9) outperforms the test in (10) in terms of BER. Yet, as Theorem 1 shows below, if we utilize the new packet-data GLRT detector in (5), we can achieve

approximately the same average[1] BER performance as with the test in (9) without the need for pure disturbance observations (secondary data) that are independent of the received data packet. The proof is given in the Appendix.

**Theorem 1** (i) *Let* $\mathbf{b}$ *be the transmitted data packet and* $\widehat{\mathbf{b}}$ *a data packet decision that differs from* $\mathbf{b}$ *in* $m$ *bits (i.e., contains* $m$ *bits in error). Then,*

$$P\left[l_1(\widehat{\mathbf{b}}) > l_1(\mathbf{b})\Big|\,\mathbf{b}\right] = \int_0^1 Q(\sqrt{2m\gamma x})\frac{x^{N-L}(1-x)^{L-2}}{B(N-L+1, L-1)}dx \tag{11}$$

*where* $Q(x) \stackrel{\triangle}{=} \frac{1}{\sqrt{2\pi}}\int_x^{+\infty} e^{-u^2/2}du$, $\gamma \stackrel{\triangle}{=} \mathbf{v}^H\mathbf{R}_n^{-1}\mathbf{v}$, *and* $l_1(\cdot)$ *is given by (6).*

(ii) *For sufficiently large transmitted energy per bit* $E$, *the average BER of the GLRT detector in (5) is approximately equal to the average BER of the scheme in (9) that utilizes* $\mathbf{R}_{n_{SA}}$ *evaluated based on* $K = N-1$ *pure disturbance observations*

$$\lim_{\gamma\to\infty} \frac{BER_{GLRT}(N)}{BER_{SMI}(N-1)} \approx 1. \tag{12}$$

■

Using Theorem 1, we can derive an approximation of the average BER of the GLRT detector and evaluate the packet size that is necessary for the GLRT detector to achieve a given BER performance level. Our findings are summarized in the following theorem whose proof is in the Appendix.

**Theorem 2** (i) *The average BER of the GLRT detector that operates on a data packet of size*

[1]The average BER of a packet-data detector is defined as the expected number of bits in error divided by the packet size. We note that the detectors in (4), (5), (9), and (10) all share the property that their average BER is equal to the BER of each bit in the packet.

*$N \geq L+2$ is given by*

$$BER_{GLRT}(N) \approx \frac{1}{\pi}\int_0^{\pi/2} M\left(N-L+1, N, -\frac{\gamma}{\sin^2\theta}\right) d\theta \tag{13}$$

$$\approx \frac{2}{3}Q\left(\sqrt{2\mu}\right)+\frac{1}{6}Q\left(\sqrt{2\mu+2\sqrt{3}\sigma}\right)+\frac{1}{6}Q\left(\sqrt{2\mu-2\sqrt{3}\sigma}\right) \tag{14}$$

*where $M(a,b,z)$ is Kummer's confluent hypergeometric function, $\gamma \triangleq \mathbf{v}^H\mathbf{R}_n^{-1}\mathbf{v}$, $\mu \triangleq \frac{N-L+1}{N}\gamma$, and $\sigma^2 \triangleq \frac{(N-L+1)(L-1)}{N^2(N+1)}\gamma^2$.*

(ii) *For any given $\nu$, the smallest packet size $N_\nu$ that guarantees that the BER performance of the GLRT packet-data detector is within $\nu$ dB from the BER performance of the optimum ML detector in (4) (i.e., $BER_{GLRT}(N) \leq Q(\sqrt{2\gamma 10^{-\nu/10}})$) is given by the ceiling of the maximum real root of the cubic equation*

$$N^3+\left(1-\frac{2(L-1)}{1-10^{-\frac{\nu}{10}}}\right)N^2+\left(\frac{(L-1)^2-3(L-1)}{\left(1-10^{-\frac{\nu}{10}}\right)^2}-\frac{2(L-1)}{1-10^{-\frac{\nu}{10}}}\right)N+\frac{4(L-1)^2}{\left(1-10^{-\frac{\nu}{10}}\right)^2}=0. \tag{15}$$

■

The roots of (15) can be obtained either numerically or analytically [13]. We note that the result in Part *(ii)* of Theorem 2 is independent of the statistics of the receiver input (i.e., independent of the performance of the optimum ML detector) which is often unknown to the designer.

### *4.1.2 Implementation Issues*

Direct implementation of (5) has complexity exponential in the number of bits. Below, we consider a procedure to obtain effective suboptimum implementations of linear complexity.

We start with an initial estimate of the packet bits $\widehat{\mathbf{b}}(0) = \left[\widehat{b}_1(0), \ldots, \widehat{b}_N(0)\right]$ replicated $P$ times to create $P$ distinct parallel search paths $\widehat{\mathbf{b}}^{(p)}(d)$, $p=1,2,\ldots,P$, $d=0,1,\ldots,D$, of final depth $D$. With each path $p$, we associate a bit index sequence $\{\pi_p(n)\}_{n=1}^N$ that is a distinct permutation of $\{1,2,\ldots,N\}$. At each step $d$, we check in parallel one bit per search path,

namely $\widehat{b}^{(p)}_{\pi_p(d \bmod N)}$ Upon completion, step $D$, we declare as our approximate GLRT decision the most likely among the $P$ bit vectors $\widehat{\mathbf{b}}^{(p)}(D)$, $p = 1, \dots, P$. The algorithm is outlined below.

Suboptimum GLRT algorithm

Initialization: Number of parallel search paths $P$; search depth $D$;

initial decision vector $\widehat{\mathbf{b}}^{(p)}(0) := \left[\widehat{b}_1(0), \widehat{b}_2(0), \dots, \widehat{b}_N(0)\right]^T, \quad p = 1, 2, \dots, P;$

search index sequences $\{\pi_p(n)\}_{n=1}^{N}$, $p = 1, 2, \dots, P$.

For step $d = 1, 2, \dots D$

For path $p = 1, 2, \dots, P$

$i := \pi_p(d \bmod N)$

$\widehat{b}_i^{(p)}(d) := \arg\max_{b_i^{(p)}} \left\{ \max_{\mathbf{R}_n} f\left(\mathbf{Y} \middle| \left\{\widehat{b}_j^{(p)}(d-1)\right\}_{j \neq i}, b_i^{(p)}, \mathbf{v}, \mathbf{R}_n\right)\right\}$

$\widehat{b}_j^{(p)}(d) := \widehat{b}_j^{(p)}(d-1), \quad j \neq i.$

end

end

$\widehat{\mathbf{b}}_{\overline{GLRT}} := \arg\max_{\mathbf{b} \in \left\{\widehat{b}^{(1)}(D), \dots, \widehat{b}^{(P)}(D)\right\}} l_1(\mathbf{b}).$

The parallel search at each step $d$ updates a single bit in each of the $P$ packet estimates (paths) and prevents early convergence to a local optimum. Using (5), the complexity of one bit update is of order $O(L)$. The overall complexity of the above algorithm for the detection of a data packet of size $N$ is of order $O(NL^2) + O(NL) + O(L^3) + O(DPL)$, which includes the cost of the initial evaluation of $\mathbf{R}_{SA}(N)$ and $\mathbf{R}_{SA}^{-1}(N)$. We note that a good initial estimate may allow relatively small values for $P$ and $D$. In this context, we can regularly reinitialize the parallel search algorithm using the best sequence estimate among the current $P$ alternatives. The performance of this *ad hoc* implementation method is examined in the simulation studies

of Section IV-C.

## 4.2 GLRT Detection: Unknown Channel

### *4.2.1 Algorithmic Development*

In this section we investigate the case of unknown channel, that is unknown $E$ and $\mathbf{a}$ according to the signal model in (2). The following proposition provides the GLRT detection scheme for this case.

**Proposition 2** *The GLRT test for the detection of the data packet* $\mathbf{b}$ *of size* $N$ *transmitted over an unknown linear channel in the presence of complex Gaussian disturbance of unknown covariance matrix* $\mathbf{R}_n$ *is given by*

$$\widehat{\mathbf{b}}_{GLRT} = \arg\max_{\mathbf{b}} \left\{ \max_{\mathbf{a},\mathbf{R}_n} f\left(\mathbf{Y} \left| \mathbf{b}, \mathbf{S}, \mathbf{a}, \mathbf{R}_n \right.\right) \right\} = \arg\max_{\mathbf{b}} \, l_2(\mathbf{b}) \tag{16}$$

$$\tag{17}$$

$$\text{where} \qquad l_2(\mathbf{b}) = \frac{N\mathbf{b}^T\mathbf{Y}^H\left[\mathbf{Y}\mathbf{Y}^H\right]^{-1}\mathbf{S}\left(\mathbf{S}^H\left[\mathbf{Y}\mathbf{Y}^H\right]^{-1}\mathbf{S}\right)^{-1}\mathbf{S}^H\left[\mathbf{Y}\mathbf{Y}^H\right]^{-1}\mathbf{Y}\mathbf{b}}{N^2 - N\mathbf{b}^T\mathbf{Y}^H\left[\mathbf{Y}\mathbf{Y}^H\right]^{-1}\mathbf{Y}\mathbf{b}}. \tag{18}$$

*Proof:* For a given bit combination $\mathbf{b}$ and channel coefficient vector $\mathbf{a}$, $f\left(\mathbf{Y} \left| \mathbf{b}, \mathbf{S}, \mathbf{a}, \mathbf{R}_n \right.\right)$ is maximized for $\mathbf{R}_n = \mathbf{R}_{n_{ML}}(\mathbf{b}, \mathbf{a}) \triangleq \frac{1}{N}\left(\mathbf{Y} - \mathbf{S}\mathbf{a}\mathbf{b}^T\right)\left(\mathbf{Y} - \mathbf{S}\mathbf{a}\mathbf{b}^T\right)^H$ and (16) becomes

$$\widehat{\mathbf{b}}_{GLRT} = \arg\max_{\mathbf{b}} \max_{\mathbf{a}} \left| \frac{1}{N}\left(\mathbf{Y} - \mathbf{S}\mathbf{a}\mathbf{b}^T\right)\left(\mathbf{Y} - \mathbf{S}\mathbf{a}\mathbf{b}^T\right)^H \right|^{-1}. \tag{19}$$

The inner maximization is solved by finding the stationary point with respect to $\mathbf{a}$. Using the identity in (8) and after some simplifications, we obtain

$$\widehat{\mathbf{a}}_{ML}(\mathbf{b}) = \frac{N\left(\mathbf{S}^H\left[\mathbf{R}_{SA}(N)\right]^{-1}\mathbf{S}\right)^{-1}\mathbf{S}^H\left[\mathbf{R}_{SA}(N)\right]^{-1}\mathbf{Y}\mathbf{b}}{N^2 - \mathbf{b}^T\mathbf{Y}^H\left[\mathbf{R}_{SA}(N)\right]^{-1}\mathbf{Y}\mathbf{b} + \mathbf{b}^T\mathbf{Y}^H\left[\mathbf{R}_{SA}(N)\right]^{-1}\mathbf{S}\left(\mathbf{S}^H\left[\mathbf{R}_{SA}(N)\right]^{-1}\mathbf{S}\right)^{-1}\mathbf{S}^H\left[\mathbf{R}_{SA}(N)\right]^{-1}\mathbf{Y}\mathbf{b}}. \tag{20}$$

Substituting (20) into (19) leads to the detection rule in (16). ■

The following theorem evaluates the asymptotic pairwise probability of error of the above detection rule in the high SNR region. The proof can be found in the Appendix.

**Theorem 3** *Let* $\mathbf{b}$ *be the transmitted data packet and* $\widehat{\mathbf{b}}$ *a data packet decision that differs from* $\mathbf{b}$ *in* $m$ *bits. Then,*

$$\lim_{E\to\infty} P\left(l_2(\widehat{\mathbf{b}}) > l_2(\mathbf{b})\middle|\,\mathbf{b}\right) = \int_0^1 Q\left(\sqrt{\frac{2m(N-m)}{N}\gamma\cdot x}\right)\frac{x^{N-L-1}(1-x)^{L-2}}{B(N-L,L-1)}dx \tag{21}$$

*where* $\gamma \triangleq \mathbf{v}^H\mathbf{R}_n^{-1}\mathbf{v}$. ■

We note that the function $l_2(\mathbf{b})$ in (18) is ambiguous with respect to the phase of $\mathbf{b}$. This phase ambiguity problem is also present in Theorem 3. In practice, phase ambiguity is resolved either by using a pilot sequence or by employing differential modulation at the transmitter; the rest of this subsection deals exactly with these two approaches.

#### *4.2.2 Pilot Assisted GLRT Detection*

**Proposition 3** *Let* $\{b_j\}_{j=1}^J$ *and* $\{b_i\}_{i=J+1}^N$ *denote, respectively,* $J$ *known pilot bits and* $N-J$ *unknown information bits within the data packet* $\mathbf{b}$ *of size* $N$ *that is transmitted over an unknown linear channel in the presence of complex Gaussian disturbance of unknown covariance. Then, the pilot assisted GLRT detector of* $\{b_i\}_{j=J+1}^N$ *is given by*

$$\left\{\widehat{b}_{i_{GLRT}}\right\}_{i=J+1}^N = \arg\max_{b_i, i\geq J+1} l_2(\mathbf{b}). \tag{22}$$

*Proof:* We note that the joint conditional pdf of the observations is given by

$$f\left(\{\mathbf{y}_i\}_{i=1}^N\middle|\{b_i\}_{i=1}^J,\{b_i\}_{i=J+1}^N,\mathbf{S},\mathbf{a},\mathbf{R}_n\right) = \frac{|\mathbf{R}_n|^{-N}}{\pi^{L\cdot N}}\exp\left\{-\sum_{i=1}^N(\mathbf{y}_i-b_i\mathbf{S}\mathbf{a})^H\mathbf{R}_n^{-1}(\mathbf{y}_i-b_i\mathbf{S}\mathbf{a})\right\}. \tag{23}$$

Thus, the GLRT detection algorithm is as in (16) with a difference only in the support of the outer optimization. ■

It is interesting to note that in (22) the pilot sequence $\{b_i\}_{i=1}^{J}$ is not used to directly estimate the phase in an explicit manner but is rather incorporated implicitly in the GLRT rule. It is also interesting to observe that the GLRT test expression in (22) maintains the same structure as in (18). For a reasonably long pilot sequence, e.g. $J \geq 2$, Theorem 3 implies that we can safely neglect the pairwise probability of error in the high SNR region for $m \geq 2$ (we note that the above pilot assisted GLRT detection rule ensures that $1 \leq m \leq N - J$ and thus eliminates the phase ambiguity problem). We conclude the treatment of the pilot assisted GLRT detection problem by deriving an approximation of the BER performance of the test in (22) and evaluating the size of the data packet that is necessary for the detector to achieve a given BER performance level (the proof utilizes Theorems 2 and 3 and is omitted due to lack of space).

**Corollary 1** (i) *The average BER of the pilot assisted GLRT detector for a data packet of size* $N \geq L + 3$ *is given by*

$$BER_{GLRT-pilot}(N) \approx BER_{SMI}(N-2) = \frac{1}{\pi}\int_0^{\pi/2} M\left(N-L, N-1, -\frac{(N-1)\gamma}{N\sin^2\theta}\right) d\theta \quad (24)$$

$$\approx \frac{2}{3}Q\left(\sqrt{2\mu}\right) + \frac{1}{6}Q\left(\sqrt{2\mu + 2\sqrt{3}\sigma}\right) + \frac{1}{6}Q\left(\sqrt{2\mu - 2\sqrt{3}\sigma}\right) \quad (25)$$

*where* $\mu \triangleq \frac{N-L}{N}\gamma$, $\sigma^2 \triangleq \frac{(N-L)(L-1)}{N^3}\gamma^2$, *and* $BER_{SMI}(N-2)$ *is the BER of the* coherent *SMI detector in (9) that would require perfect knowledge of* $\mathbf{a}$ *and utilize* $K = N - 2$ *independent pure disturbance observations.*

(ii) *For any given* $\nu$, *the smallest packet size* $N_\nu$ *that guarantees that the BER performance of the GLRT pilot assisted packet-data detector is within* $\nu$ *dB from the BER performance of the*

*optimum* coherent *ML detector in (4) (i.e.,* $BER_{GLRT-pilot}(N) \leq Q(\sqrt{2\gamma 10^{-\nu/10}})$*) is given by the ceiling of the maximum real root of the cubic equation*

$$N^3 - \frac{2L}{1-10^{-\frac{\nu}{10}}}N^2 + \frac{L^2-3(L-1)}{\left(1-10^{-\frac{\nu}{10}}\right)^2}N + \frac{3L(L-1)}{\left(1-10^{-\frac{\nu}{10}}\right)^2} = 0. \tag{26}$$

■

Corollary 1 implies that the pilot assisted GLRT detector performs closely to the *coherent* SMI detector in (9) that requires perfect knowledge of $\mathbf{a}$ and assumes availability of pure disturbance observations. We can modify in a straightforward manner the algorithm outlined in Section III-A to obtain a suboptimum implementation of the unknown-channel pilot GLRT scheme in (22) that exhibits linear complexity.

### *4.2.3 DPSK GLRT Detection*

As an alternative to pilot signaling, phase ambiguity of the GLRT detector in (18) can be resolved by employing differential encoding at the transmitter. To avoid redundancy in our presentation, in this section we keep the size of the transmitted packet equal to $N$ while the number of the information bits embedded in the differentially encoded packet is $N-1$, $\{b_i\}_{i=1}^{N-1}$. The differentially encoded bits themselves are $d_0 = +1$ and $d_i = d_{i-1}b_i$, $i = 1, 2, \ldots, N-1$. The $i$th received vector $\mathbf{y}_i$ is still of the form of (2) with $d_i$ in place of $b_i$. Given the transmitted bits $d_i$, $i = 0, 1 \ldots N-1$, the information bits can be uniquely determined by $b_i = d_{i-1}d_i$, $i = 1, \ldots, N-1$.

We recall [14] that under ideal conditions (i.e., perfectly known interference-plus-noise statistics and channel impulse response) the ideal optimum (ML) differential detector of a packet of $N-1$ information bits consists of the ideal linear filter $\mathbf{R}_n^{-1}\mathbf{S}\mathbf{a}$ followed first by a sign detector and then by the 2-symbol block differential decoder. On the other hand, when interference-plus-

noise statistics are perfectly known but the channel is known only within a phase ambiguity $\theta$, the optimum (ML) differential detector of a block (packet) of $N-1$ information bits consists of the ideal linear filter $\mathbf{R}_n^{-1}\mathbf{S}\mathbf{a}e^{j\theta}$ followed by an $N$-symbol differential decoder that operates on a block of complex, in general, scalar outputs of the optimum linear filter [15] (the linear filter $\mathbf{R}_n^{-1}\mathbf{S}\mathbf{a}e^{j\theta}$ provides the sufficient statistics for differential decoding). Direct implementation of the optimum block differential decoder according to the likelihood metric requires exponential complexity in the block (packet) size $N$ (fast approximate algorithms with complexity $N\log N$ can be used instead [16]- [19]). It is well understood that the BER performance of the phase-ambiguity-optimum block differential detector is lower bounded by the BER performance of the ideal all-known differential detector and approaches this lower bound as $N \to \infty$. A popular suboptimum receiver for differentially encoded data has been the 2-symbol differential detector that utilizes a 2-symbol only differential decoder and detects one bit at a time by evaluating the sign of the real part of the product of the current filter output with the previous conjugated filter output.

Under unknown input statistics and channel coefficients, the common approach has been to produce estimates of the unknown quantities and insert the estimates in the N-symbol (or 2-symbol) block differential detector. Instead, what we propose herein is a GLRT-type scheme that combines into a single optimization effort estimation of interference-plus-noise covariance matrix and channel coefficients and detection of packet data. The following proposition identifies our DPSK GLRT scheme.

**Proposition 4** *The DPSK GLRT detector of differentially encoded packet data* $\{b_i\}_{i=1}^{N-1}$ *transmitted over an unknown linear channel in the presence of complex Gaussian disturbance of*

*unknown covariance is given by*

$$\left\{\widehat{d}_{i_{GLRT}}\right\}_{i=1}^{N-1} = \arg\max_{d_i, i\geq 1} l_2(\mathbf{d}) \tag{27}$$

$$\widehat{b}_{i_{GLRT}} = \widehat{d}_{i-1_{GLRT}}\widehat{d}_{i_{GLRT}}, \quad i = 1, 2, \ldots, N-1, \tag{28}$$

*where* $\mathbf{d} \stackrel{\triangle}{=} [d_0, \ldots d_{N-1}]^T$.

*Proof:* It suffices to observe that the one-to-one mapping of information bits to differentially encoded bits implies that maximization of the generalized likelihood function with respect to the information bits is equivalent to maximization with respect to the differentially encoded (transmitted) bits. ■

Using Theorem 3 and the observation that both a single-bit error and an $(N-1)$-bit error in $\widehat{\mathbf{d}}$ results in a 2-bit error in the bit sequence $\left\{\widehat{b}_i\right\}_{i=1}^{N-1}$, we can approximate the BER performance of the DPSK GLRT detector and obtain the packet size that is necessary for the detector to achieve a certain BER performance level as follows.

**Corollary 2** (i) *The average BER of the DPSK GLRT detector that operates on a data packet of size* $N \geq L+3$ *is given by*

$$BER_{GLRT-DPSK}(N) \approx BER_{SMI,DPSK}(N-2) \approx \frac{2}{\pi}\int_0^{\pi/2} M\left(N-L, N-1, -\frac{(N-1)\gamma}{N\sin^2\theta}\right) d\theta \tag{29}$$

$$\approx \frac{4}{3}Q\left(\sqrt{2\mu}\right) + \frac{1}{3}Q\left(\sqrt{2\mu + 2\sqrt{3}\sigma}\right) + \frac{1}{3}Q\left(\sqrt{2\mu - 2\sqrt{3}\sigma}\right) \tag{30}$$

*where* $\mu \stackrel{\triangle}{=} \frac{N-L}{N}\gamma$, $\sigma^2 \stackrel{\triangle}{=} \frac{(N-L)(L-1)}{N^3}\gamma^2$, *and* $BER_{SMI,DPSK}(N-2)$ *is the BER of a detection scheme that would utilize the* coherent *SMI detector of* $\{d_i\}_{i=1}^N$ *built on* $K = N-2$ *independent pure disturbance observations under perfect knowledge of* $\mathbf{a}$*, followed by the differential decoder* $\widehat{b}_i = \widehat{d}_{i-1}\widehat{d}_i$, $i = 1, \ldots, N-1$.

(ii) *For any given $\nu$, the smallest packet size $N_\nu$ that guarantees that the BER performance of the DPSK GLRT packet-data detector is within $\nu$ dB from the BER performance of the optimum ML DPSK detector (i.e., $BER_{GLRT,DPSK}(N) \leq 2Q(\sqrt{2\gamma 10^{-\nu/10}})$) is given by the ceiling of the maximum real root of the cubic equation*

$$N^3 - \frac{2L}{1-10^{-\frac{\nu}{10}}}N^2 + \frac{L^2-3(L-1)}{\left(1-10^{-\frac{\nu}{10}}\right)^2}N + \frac{3L(L-1)}{\left(1-10^{-\frac{\nu}{10}}\right)^2} = 0. \tag{31}$$

■

In the following section, simulation studies demonstrate that the proposed DPSK GLRT detector that combines estimation of the unknown parameters and packet-data detection into one process outperforms the common estimate-and-plug-in approach where we first take the optimum $N$-symbol (or the popular suboptimum 2-symbol) block differential detector formula and then substitute therein unknown statistics and channel coefficients by estimates obtained separately.

## 4.3 Simulation Studies

We prepare a communication system simulation study where packet-data are received in the presence of Gaussian noise of unknown covariance. The covariance matrix used to generate received data is taken directly from the literature [20]. The dimension of the received data vectors is $L = 9$. The channel processed signal waveform $\mathbf{Sa}$ is chosen arbitrarily and is assumed to be known (known channel case). We would like to study the BER performance of the proposed GLRT packet-data detector as a function of the data packet size $N$ and examine the accuracy of our analytical BER expression in (14). The GLRT detector is implemented in its linear cost form as presented in Section IV.A with $P = 16$, $D = 6N$, and arbitrary initial bit

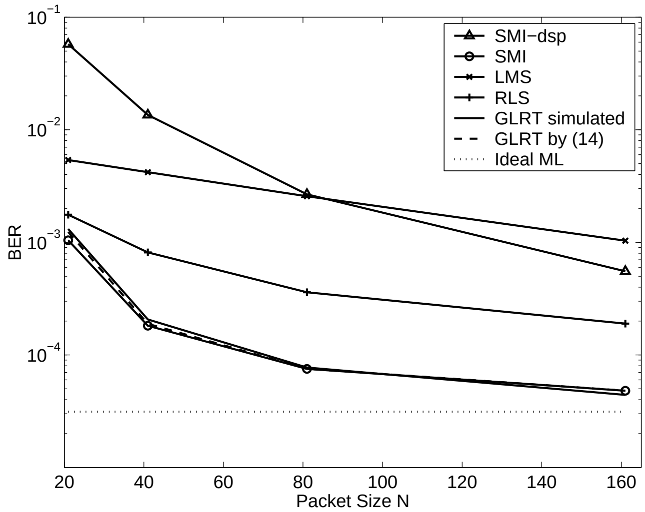


Fig. 1. BER of packet-data detectors as a function of the packet size $N$ (Gaussian disturbance of unknown covariance, known channel, $E = 7dB$).

estimates. Fig. 1 presents our findings and comparisons with SMI-dsp (desired-signal-present) in (10), SMI (disturbance only observations) in (9), LMS (step size $10^{-3}$), RLS (initialization parameter 100), and ideal ML detection. In view of the nearly overlapping analytical and simulated GLRT BER curves, we may claim that our linear cost GLRT implementation performs very close to full GLRT and (14) provides an accurate approximation of the BER of the GLRT packet-data detector. Furthermore, the GLRT packet-data detector outperforms significantly the SMI-dsp, LMS, and RLS detectors and performs nearly the same as the SMI detector in (9) that requires $N-1$ *additional* pure disturbance observations.

In the rest of this section, we use as an illustration vehicle for the proposed GLRT schemes a packet-data DS-CDMA communication system[2]. At all times, the GLRT detectors are imple-

[2]The combined effect of DS-CDMA multiple access interference (MAI) and AWGN is Gaussian-mixture distributed and not plain Gaussian. It is interesting to examine how the newly developed GLRT detectors

mented via the linear complexity algorithm of Section IV.B with $P = 16$ and $D = 6N$. Initial bit estimates are taken either by conventional matched-filter (MF) outputs (Case-study #1) or are arbitrarily set (Case-studies #2 and #3).

*DS-CDMA Case-study #1 Synchronous multiuser system and single-path channel*

We consider a system with 10 synchronous users with Gold signatures of length $G = 31$. We select a "user of interest" and have the SNR's of the interfering users fixed in the range $[6dB, 11dB]$. In this study we assume exact knowledge of the channel of the user of interest. We compare the BER of the GLRT detector with the BER of the MF, SMI-dsp, LMS (step size $10^{-4}$), RLS (initialization parameter 100), and SMI detector in (9) that assumes availability of $N - 1$ additional pure disturbance observations.

In Figs. 2 and 3, we plot the BER as a function of the SNR of the user of interest and the packet size $N$, respectively. We observe that the GLRT detector performs very closely to the SMI detector in (9) and outperforms all other detectors.

*DS-CDMA Case-study #2 Asynchronous multipath fading channel: Pilot-assisted signaling*

We consider the same setup as in Case-study #1, except that now the users transmit asynchronously and each user channel has 3 resolvable paths. The path coefficients are modeled as independent complex Gaussian random variables all of unit variance. The length of the pilot sequence is fixed at $J = 10$. We compare the BER of the GLRT detector with the BER of the RAKE-MF, the SMI and SMI-dsp detectors in (9) and (10), and the LMS and RLS detectors. We note that in this study the GLRT detector assumes no knowledge of the channel *while all other detectors assume exact knowledge of the channel.* In addition, the SMI detector in (9) uses $N - 2$ *extra* pure disturbance observations that are assumed to be available. It is

perform in such an environment.

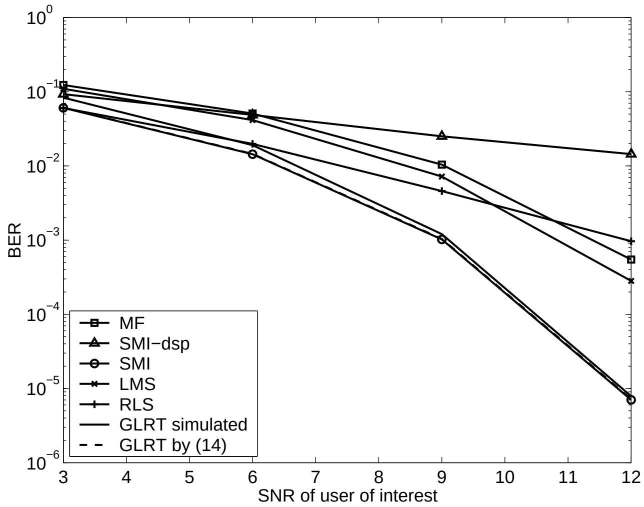


Fig. 2. Case-study #1: BER as function of the SNR of the user of interest ($N = 127$).

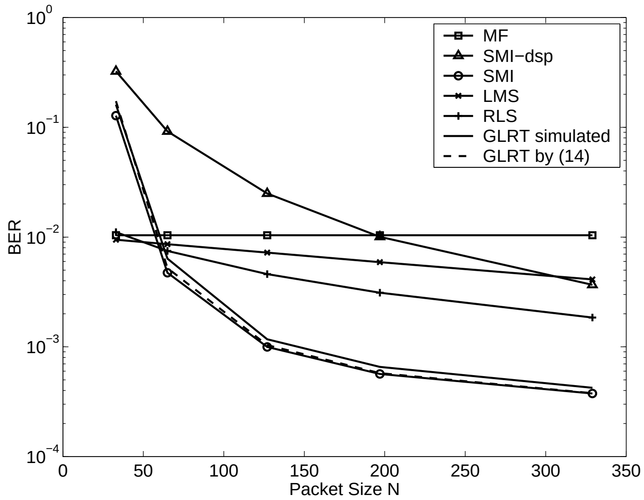


Fig. 3. Case-study #1: BER as function of the packet size $N$. The SNR of the user of interest is fixed at $9dB$.

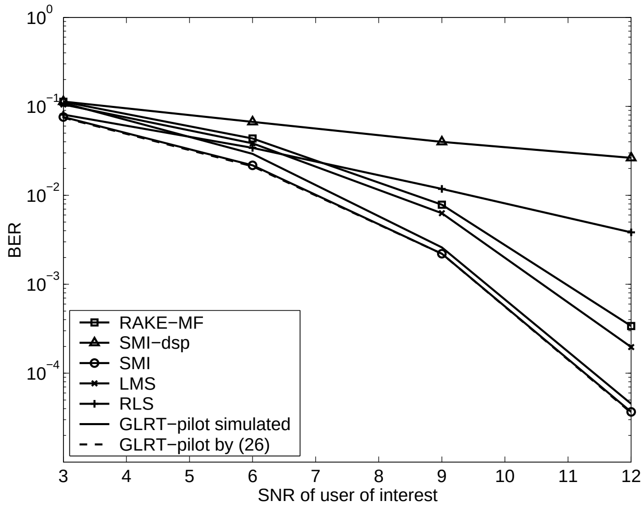


Fig. 4. Case-study #2: BER as function of the SNR of the user of interest ($N = 127$).

worth noting that the pilot sequence is incorporated and processed internally and elegantly by the GLRT algorithm without the need for a separate phase estimation stage. Our simulation findings given in Figs. 4 and 5 are self-explanatory and make a strong case in favor of the new GLRT developments.

*DS-CDMA Case-study #3 Asynchronous multipath fading channel: DPSK signaling*

We consider the same setup as in Case-study #2, except that the transmitter now uses DPSK encoding instead of pilot signaling; hence, at the receiver end a differential decoder is needed to recover the information bits. We compare the BER of our DPSK GLRT detector with the BER of the DPSK version of the following detectors: RAKE-MF, SMI-dsp, LMS, RLS, and ideal MMSE. A 2-symbol differential decoder is used in all cases. The *coherent* SMI detector that is described in Corollary 2, Part *(i)* is also included as a reference. We note that the DPSK GLRT detector assumes no knowledge of the channel while the RAKE-MF, LMS,

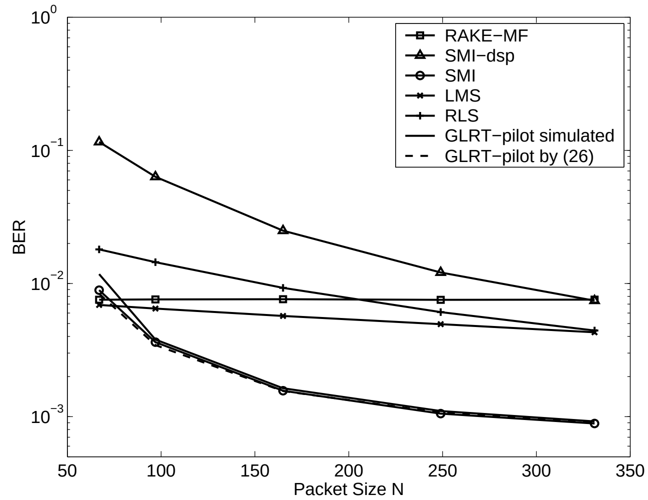


Fig. 5. Case-study #2: BER as function of the packet size $N$. The SNR of the user of interest is fixed at $9dB$.

RLS, SMI-dsp, and ideal MMSE detectors assume perfectly known path coefficients *up to an unknown phase* (phase ambiguity is resolved by differential encoding/decoding). In addition, the ideal MMSE detector assumes perfectly known interference-plus-noise covariance matrix and the coherent SMI detector in Corollary 2, Part *(i)* requires perfect knowledge of the path coefficients (including the phase) and $N-2$ additional pure disturbance observations. In Figs. 6 and 7 we repeat the studies of Figs. 4 and 5. The superiority of the new DPSK GLRT detector is striking. In fact, for data packets of size $N = 250$ and higher the DPSK GLRT detector outperforms even the ideal MMSE (2-symbol decoder) detector.

# 5. Conclusions

We considered the problem of packet-data detection for general multiaccess/multiuser digital communication systems. We proposed novel GLRT-type detection schemes that perform joint

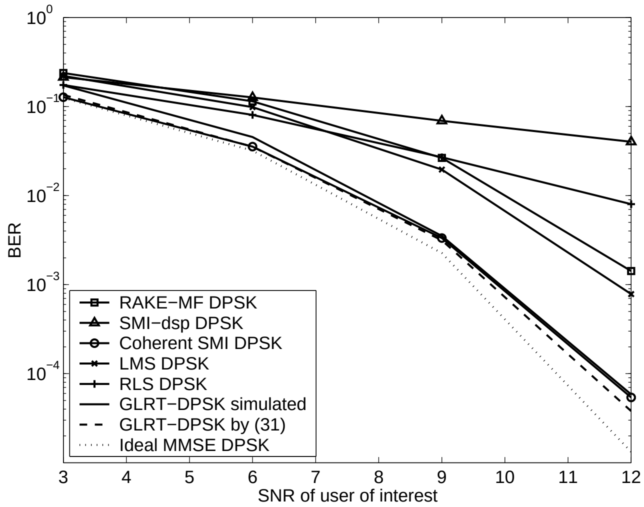


Fig. 6. Case-study #3: BER as function of the SNR of the user of interest ($N = 160$).

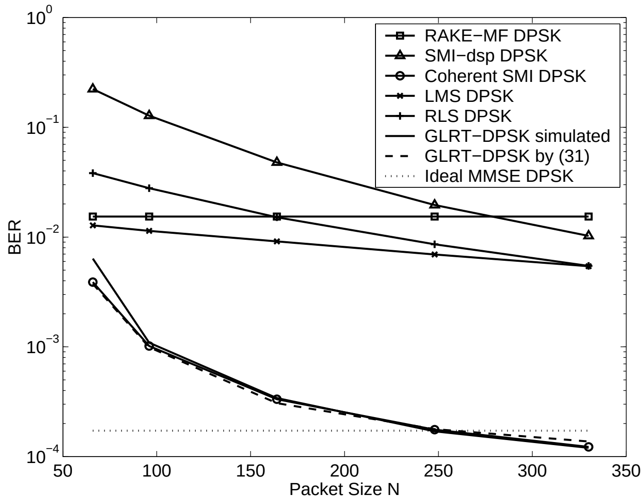


Fig. 7. Case-study #3: BER as function of the packet size $N$. The SNR of the user of interest is fixed at $9dB$.

estimation of the unknown system parameters and detection of the packet-data and we designed suboptimum implementations of linear complexity in the packet size. In particular, for the known channel case we developed a coherent GLRT detector, while for the unknown channel case we developed a pilot assisted GLRT detector (the pilot signal is implicitly used to resolve phase ambiguity) and a DPSK GLRT detector. We established analytically the performance of each proposed GLRT-type scheme relative to the corresponding conventional estimate-and-plug-in detector that replaces unknown parameters in its ideal formula by estimates obtained through a separate estimation process. In all cases, the GLRT schemes maintain the same elegant core structure regardless of known or unknown channels and pilot or DPSK signaling. Finally, we developed analytical performance evaluation tools that provide the BER that can be reached by each proposed GLRT scheme for a given finite data record size, as well as the data record size that is necessary for each GLRT detector to perform within a certain neighborhood of the optimal performance point (*without the need to know the latter*).

# 6. References


[1] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part I.* New York: Wiley, 1968.

[2] I. Motedayen-Aval and A. Anastasopoulos, "Polynomial-complexity noncoherent symbol-by-symbol detection with application to adaptive iterative decoding of turbo-like codes," *IEEE Trans. Commun.,* vol. 51, pp. 197–207, Feb. 2003.

[3] E. J. Kelly, "An adaptive detection algorithm," *IEEE Trans. Aerosp. Electron. Syst.,* vol. AES-22, pp. 115-127, Jan. 1986.

[4] F. C. Robey, D. R. Fuhrmann, E. J. Kelly, and R. Nitzberg, "A CFAR adaptive matched filter detector," *IEEE Trans. Aerosp. Electron. Syst.,* vol. 28, pp. 208-216, Jan. 1992.

[5] S. M. Kay, V. Nagesha, and J. Salisbury, "Broad-band detection based on two-dimensional mixed autoregressive models," *IEEE Trans. Signal Proc.,* pp. 2413–2428, July 1993.

[6] S. Kraut, L. L. Scharf, and L. T. McWhorter, "Adaptive subspace detectors," *IEEE Trans. Signal Proc.,* vol. 49, pp. 1-16, Jan. 2001.

[7] M. K. Varanasi, "Group detection for synchronous Gaussian code-division multiple-access channels," *IEEE Trans. Info. Theory,* vol. 41, pp. 1083–1096, July 1995.

[8] M. K. Varanasi, "Blind multiuser detection via interference identification," *IEEE Trans. Commun.,* vol. 50, pp. 1172–1181, July 2002.

[9] I. N. Psaromiligkos and S. N. Batalama, "Data record size requirements for adaptive space-time DS/CDMA signal detection and direction of arrival estimation," *IEEE Trans. Commun.*, submitted.

[10] I. N. Psaromiligkos and S. N. Batalama, "Data record size requirements for adaptive antenna arrays," in *Proc. SPIE, Dig. Wireless Comm. Conf.*, Orlando, FL, April 2000, vol. 4045, pp. 122-131.

[11] I. N. Psaromiligkos and S. N. Batalama, "Recursive AV and MVDR filter estimation for maximum SINR adaptive space-time processing," *IEEE Trans. Commun.*, to appear June 2004.

[12] J. M. Farrell, I. N. Psaromiligkos, and S. N. Batalama, "Design and analysis of supervised and decision-directed estimators of the MMSE/LCMV filter in data-limited environments," in *Proc. SPIE, Dig. Wireless. Comm. Conf.*, Orlando, FL, 2003, vol. 5100, pp. 227-237.

[13] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions.* New York, NY: Dover, 1970.

[14] M. K. Simon, S. M. Hinedi, and W. C. Lindsey, *Digital Communication Techniques: Signal Design and Detection.* Upper Saddle River, NJ: Prentice Hall, 1995.

[15] D. Divsalar and M. K. Simon, "Multiple-symbol differential detection of MPSK," *IEEE Trans. Commun.,* vol. 38, pp. 300–308, Mar. 1990.

[16] H. Leib and S. Pasupathy, "The phase of a vector perturbed by Gaussian noise and differentially coherent receivers," *IEEE Trans. Info. Theory,* vol. 34, pp. 1491–1500, Nov. 1988.

[17] H. Leib, "Data-aided noncoherent demodulation of DPSK," *IEEE Trans. Commun.,* vol. 43, pp. 722–725, Feb./Mar./Apr. 1995.

[18] K. M. Machenthun, "A fast algorithm for multiple-symbol differential detection of MPSK," *IEEE Trans. Commun.,* vol. 42, pp. 1471–1474, Feb./Mar./Apr. 1994.

[19] W. Sweldens, "Fast block noncoherent decoding," *IEEE Commun. Lett.,* vol. 5, pp. 132–134, Apr. 2001.

[20] D. B. Rubin and D. T. Thayer, "EM algorithms for ML factor analysis," *Psychometrika,* vol. 47, pp. 69-76, 1982.

[21] R. A. Horn and C. R. Johnson, *Matrix Analysis.* Cambridge, UK: Cambridge University Press, 1985.

[22] R. J. Muirhead, *Aspects of Multivariate Statistical Theory.* New York: Wiley, 1982.

[23] I. S. Reed, J. D. Mallett, and L. E. Brennan, "Rapid convergence rate in adaptive arrays," *IEEE Trans. Aerosp. Electron. Syst.,* vol. 10, pp. 853-863, Nov. 1974.

[24] M. K. Simon and M. Alouini, *Digital Communication over Fading Channels.* New York: Wiley, 2000.

[25] J. M. Holtzman, "A simple, accurate method to calculate spread-spectrum multiple-access error probabilities," *IEEE Trans. Commun.,* vol. 40, pp. 461-464, Mar. 1992.

# 7. Appendices

**Proof of Theorem 1**

*(i)* We can write $\mathbf{Y} = \mathbf{v}\mathbf{b}^T + \mathbf{N}$ where $\mathbf{N} \triangleq [\mathbf{n}_1, \dots, \mathbf{n}_N] \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_N \otimes \mathbf{R}_n)$. W.l.o.g. assume $\mathbf{b} = [\underbrace{1, \dots, 1}_{N}]^T$ and $\widehat{\mathbf{b}} = [\underbrace{-1, \dots, -1}_{m}, 1, \dots, 1]^T$. By (5) and (7), $l_1(\widehat{\mathbf{b}}) > l_1(\mathbf{b})$ is equivalent to

$$\left|\mathbf{N}\mathbf{N}^H + 2\mathbf{v}\left(\sum_{i=1}^{m} \mathbf{n}_i\right)^H + 2\left(\sum_{i=1}^{m} \mathbf{n}_i\right)\mathbf{v}^H + 4m\mathbf{v}\mathbf{v}^H\right| < \left|\mathbf{N}\mathbf{N}^H\right|. \tag{32}$$

Let $\mathbf{U} \triangleq [\mathbf{u}_1, \dots, \mathbf{u}_N]$ be a unitary matrix and define $\mathbf{N}' \triangleq \mathbf{N}\mathbf{U} = [\mathbf{n}_1', \mathbf{N}'']$ where $\mathbf{n}_1' = \mathbf{N}\mathbf{u}_1$ and $\mathbf{u}_1 = [\underbrace{1, \dots, 1}_{m}, 0, \dots, 0]^T/\sqrt{m}$. Then, $\mathbf{N}' \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_N \otimes \mathbf{R}_n)$, $\mathbf{n}_1' = \frac{1}{\sqrt{m}} \sum_{i=1}^{m} \mathbf{n}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_n)$,

$\mathbf{N}'' \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_{N-1} \otimes \mathbf{R}_n)$, and $\mathbf{N}\mathbf{N}^H = \mathbf{N}'\mathbf{N}'^H = \mathbf{N}''\mathbf{N}''^H + \mathbf{n}_1'\mathbf{n}_1'^H$ Thus, (32) can be reduced to

$$\left|\mathbf{A} + \left(\mathbf{n}_1' + 2\sqrt{m}\mathbf{v}\right)\left(\mathbf{n}_1' + 2\sqrt{m}\mathbf{v}\right)^H\right| < \left|\mathbf{A} + \mathbf{n}_1'\mathbf{n}_1'^H\right| \tag{33}$$

where $\mathbf{A} \triangleq \mathbf{N}''\mathbf{N}''^H \sim \mathcal{C}W_L(N-1, \mathbf{R}_n)$ is complex Wishart distributed with $N-1$ degrees of freedom and independent of $\mathbf{n}_1'$ [22]. Using identity (8), we simplify (33) to

$$\mathrm{Re}\left(\mathbf{v}^H\mathbf{A}^{-1}\mathbf{n}_1'\right) < -\sqrt{m}\left(\mathbf{v}^H\mathbf{A}^{-1}\mathbf{v}\right). \tag{34}$$

Given $\mathbf{A}$, $\mathbf{v}^H\mathbf{A}^{-1}\mathbf{n}_1'$ is circular complex Gaussian distributed with zero mean and variance $\mathbf{v}^H\mathbf{A}^{-1}\mathbf{R}_n\mathbf{A}^{-1}\mathbf{v}$. Thus, given $\mathbf{A}$ and $\mathbf{b}$

$$P\left(l_1(\widehat{\mathbf{b}}) > l_1(\mathbf{b})\middle|\, \mathbf{A}, \mathbf{b}\right) = Q\left(\sqrt{\frac{2m\left(\mathbf{v}^H\mathbf{A}^{-1}\mathbf{v}\right)^2}{\mathbf{v}^H\mathbf{A}^{-1}\mathbf{R}_n\mathbf{A}^{-1}\mathbf{v}}}\right). \tag{35}$$

Using a similar reasoning as in [23], $\eta \triangleq \frac{\left(\mathbf{v}^H\mathbf{A}^{-1}\mathbf{v}\right)^2}{\gamma(\mathbf{v}^H\mathbf{A}^{-1}\mathbf{R}_n\mathbf{A}^{-1}\mathbf{v})}$ has a Beta distribution $f_\eta(x) = \frac{x^{a-1}(1-x)^{b-1}}{B(a,b)}$, $0 \le x \le 1$, where $a = N-L+1$, $b = L-1$, and $B(a,b)$ is the Beta function [13]. Then, (11) is just the expected value of (35) with respect to $\mathbf{A}$ (or equivalently $\eta$).

*(ii)* For $m \ge 2$ we can show that

$$\lim_{\gamma\to\infty}\frac{\int_0^1 Q\left(\sqrt{2m\gamma x}\right)f_\eta(x)dx}{\int_0^1 Q\left(\sqrt{2\gamma x}\right)f_\eta(x)dx} = \lim_{\gamma\to\infty}\frac{\sqrt{m}\frac{\Gamma(a+b+0.5)}{\Gamma(b)}(m\gamma)^{-a-0.5}\left(1+O(|\gamma|^{-1})\right)}{\frac{\Gamma(a+b+0.5)}{\Gamma(b)}(\gamma)^{-a-0.5}\left(1+O(|\gamma|^{-1})\right)} = m^{-(N-L+1.5)} \tag{36}$$

where $\Gamma(\cdot)$ is the complete Gamma function. The sequence of operations that lead to this result in (36) includes application of L'Hospital's rule, use of the fact that the moment generating function of a Beta distributed variable with parameters $a$ and $b$ is a Kummer's confluent hypergeometric function [13], i.e., $\int_0^\infty e^{sx}f_\rho(x)dx = M(a, a+b, s)$, and use of the asymptotic expression of $M(a, a+b, s)$ [13]. (36) implies that the pairwise probability of error in (11) is

negligible in the high received SNR region for $m \geq 2$. Under this approximation,

$$BER_{GLRT}(N) = \frac{1}{N}\sum_{\widehat{\mathbf{b}}} P\left(\widehat{\mathbf{b}}_{GLRT} = \widehat{\mathbf{b}} \,\middle|\, \mathbf{b}\right) m(\widehat{\mathbf{b}}) \leq \frac{1}{N}\sum_{\widehat{\mathbf{b}}} P\left(l_1(\widehat{\mathbf{b}}) > l_1(\mathbf{b}) \,\middle|\, \mathbf{b}\right) m(\widehat{\mathbf{b}}) \tag{37}$$

$$\approx \frac{1}{N}\sum_{m(\widehat{\mathbf{b}})=1} \int_0^1 Q\left(\sqrt{2\gamma x}\right) f_\eta(x)dx = \int_0^1 Q\left(\sqrt{2\gamma x}\right) f_\eta(x)dx \tag{38}$$

where $m(\widehat{\mathbf{b}})$ is the number of bits that $\widehat{\mathbf{b}}$ differs from $\mathbf{b}$. Next, we consider the average BER performance of the SMI detector in (9). Define

$$\rho \triangleq \frac{\left|\mathbf{v}^H \mathbf{R}_{nSA}^{-1}(N-1)\mathbf{v}\right|^2}{\gamma\left(\mathbf{v}^H \mathbf{R}_{nSA}^{-1}(N-1)\mathbf{R}_n \mathbf{R}_{nSA}^{-1}(N-1)\mathbf{v}\right)} \tag{39}$$

where $\mathbf{R}_{nSA}(N-1)$ is evaluated based on $N-1$ pure disturbance observations[3]. By [23], $\rho$ is Beta distributed with parameters $a = N - L + 1$ and $b = L - 1$ (hence $\rho$ and $\eta$ have identical pdf's). The average BER of the detector in (9) is

$$BER_{SMI}(N-1) = \int_0^1 Q\left(\sqrt{2\gamma x}\right) f_\rho(x)dx = \int_0^1 Q\left(\sqrt{2\gamma x}\right) f_\eta(x)dx \tag{40}$$

Expressions (38) and (40) provide an upper bound on $BER_{GLRT}(N)$. On the other hand, the average BER of the GLRT packet-data detector is equal to the BER of any bit, say bit $b_N$. The BER of detecting $b_N$ using the GLRT packet-data detector in (5) is lower bounded by the BER of the following GLRT-type detection scheme that assumes perfect knowledge of the first $N-1$ bits in the packet:

$$\widehat{b}_N = \arg\max_{b_N} \left\{ \max_{\mathbf{R}_n} f\left(\mathbf{Y} \,\middle|\, \{b_i\}_{i=1}^{N-1}, b_N, \mathbf{v}, \mathbf{R}_n\right) \right\}. \tag{41}$$

Following similar calculations used to obtain (7) and then applying (8), we have

$$\widehat{b}_N = \arg\max_{b_N} \left[\mathrm{Re}\left(\mathbf{v}^H \mathbf{R}_{nSA}^{-1}(N-1)\mathbf{y}_N b_N\right)\right] = \mathrm{sgn}\left[\mathrm{Re}\left(\mathbf{v}^H \mathbf{R}_{nSA}^{-1}(N-1)\mathbf{y}_N\right)\right] \tag{42}$$

[3]We recall that $N-1 \geq L$ is the necessary condition to guarantee the existence of the inverse of $\mathbf{R}_{nSA}(N-1)$ with probability 1.

where $\mathbf{R}_{n_{SA}}(N-1)$ is the sample average noise correlation matrix evaluated using the first $N-1$ noise components. We see that (42) is the detector in (9) with $K = N-1$. Hence, the BER of the detection scheme in (41) is equal to $BER_{SMI}(N-1)$, and

$$BER_{GLRT}(N) \geq BER_{SMI}(N-1). \tag{43}$$

If we combine (38), (40) and (43), then (12) follows. ■

**Proof of Theorem 2**

*(i)* Expression (13) can be proved by virtue of Theorem 1, expression (5.3) of [24], and the fact that the moment generating function of a Beta distributed random variable is a Kummer's confluent hypergeometric function [13]. Alternatively, we can express $BER_{GLRT}(N)$ as $E_\theta\{g(\theta)\}$ for $g(\theta) \triangleq Q(\sqrt{2\theta})$, $\theta \triangleq \gamma\eta$ and $\eta$ a Beta distributed random variable with $a = N-L+1$ and $b = L-1$ [13], and then approximate by $\frac{2}{3}g(\mu) + \frac{1}{6}g(\mu+\sqrt{3}\sigma) + \frac{1}{6}g(\mu-\sqrt{3}\sigma)$ [25] for $\mu \triangleq \frac{N-L+1}{N}\gamma$, $\sigma^2 \triangleq \frac{(N-L+1)(L-1)}{N^2(N+1)}\gamma^2$, and any $\frac{N-L+1}{N} \geq \sqrt{\frac{3(N-L+1)(L-1)}{N^2(N+1)}}$. The latter inequality holds for all $L$ and $N > L+2$ which also satisfies the condition for the existence of the inverse of $\mathbf{R}_{n_{SA}}(N-1)$ with probability 1.

*(ii)* Let $h(N) \triangleq \frac{N-L+1}{N} - \sqrt{\frac{3(N-L+1)(L-1)}{N^2(N+1)}}$. The monotonicity of $Q(x)$ and (14) imply that

$$\frac{1}{6}Q\left(\sqrt{2\gamma h(N)}\right) < BER_{GLRT}(N) < Q\left(\sqrt{2\gamma h(N)}\right). \tag{44}$$

Thus, $Q\left(\sqrt{2\gamma h(N)}\right)$ is an asymptotically tight upper bound on $BER_{GLRT}(N)$. The smallest $N_\nu$ that guarantees $BER_{GLRT}(N) \leq Q(\sqrt{2\gamma 10^{-\nu/10}})$ $\forall\gamma$ is the ceiling of the solution of the equation $h(N) = 10^{-\nu/10}$ which can be found as the maximum real root of (15). ■

**Proof of Theorem 3**

W.l.o.g. assume $\mathbf{b} = [\underbrace{1,\ldots,1}_{N}]$ and $\widehat{\mathbf{b}} = [\underbrace{-1,\ldots,-1}_{m},1,\ldots,1]$. Let $\overline{\mathbf{y}}_1 \triangleq \frac{1}{m}\sum_{i=1}^{m}\mathbf{y}_i$, $\overline{\mathbf{y}}_2 \triangleq$

$\frac{1}{N-m}\sum_{i=m+1}^{N}\mathbf{y}_i$, $\mathbf{A}_1 \triangleq \sum_{i=1}^{m}(\mathbf{y}_i-\overline{\mathbf{y}}_1)(\mathbf{y}_i-\overline{\mathbf{y}}_1)^H$, $\mathbf{A}_2 \triangleq \sum_{i=m+1}^{N}(\mathbf{y}_i-\overline{\mathbf{y}}_2)(\mathbf{y}_i-\overline{\mathbf{y}}_1)^H$, and $\mathbf{A}=\mathbf{A}_1+\mathbf{A}_2$. Then $\overline{\mathbf{y}}_1 \sim N(\mathbf{v},\mathbf{R}_n/m)$ is independent of $\mathbf{A}_1$ and $\overline{\mathbf{y}}_2 \sim N(\mathbf{v},\mathbf{R}_n/(N-m))$ is independent of $\mathbf{A}_2$ [22]. We also note that $\overline{\mathbf{y}}_1$ is independent of $\mathbf{A}_2$ and $\overline{\mathbf{y}}_2$ is independent of $\mathbf{A}_1$. Thus, $\overline{\mathbf{Y}} \triangleq [\overline{\mathbf{y}}_1,\overline{\mathbf{y}}_2]$ is independent of $\mathbf{A}$, which has a complex Wishart distribution with $N-2$ degrees of freedom [22]. Set $\mathbf{D} \triangleq \operatorname{diag}(m,N-m)$. Then $\mathbf{Y}\mathbf{Y}^H=\mathbf{A}+m\overline{\mathbf{y}}_1\overline{\mathbf{y}}_1^H+(N-m)\overline{\mathbf{y}}_2\overline{\mathbf{y}}_2^H=\mathbf{A}+\overline{\mathbf{Y}}\mathbf{D}\overline{\mathbf{Y}}^H$. Through straightforward -yet tedious- matrix algebra [21], we have

$$(\mathbf{Y}\mathbf{Y}^H)^{-1}=\mathbf{A}^{-1}-\mathbf{A}^{-1}\overline{\mathbf{Y}}\left(\mathbf{D}^{-1}+\overline{\mathbf{Y}}^H\mathbf{A}^{-1}\overline{\mathbf{Y}}\right)^{-1}\overline{\mathbf{Y}}^H\mathbf{A}^{-1}, \tag{45}$$

$$\left[\mathbf{S}^H(\mathbf{Y}\mathbf{Y}^H)^{-1}\mathbf{S}\right]^{-1}=\mathbf{B}^{-1}+\mathbf{B}^{-1}\mathbf{Z}\left[\mathbf{D}^{-1}+\overline{\mathbf{Y}}^H\mathbf{A}^{-1}\overline{\mathbf{Y}}-\mathbf{Z}^H\mathbf{B}^{-1}\mathbf{Z}\right]^{-1}\mathbf{Z}^H\mathbf{B}^{-1}, \tag{46}$$

$$\mathbf{D}^{-1}-\overline{\mathbf{Y}}^H(\mathbf{Y}\mathbf{Y}^H)^{-1}\overline{\mathbf{Y}}=\mathbf{D}^{-1}\left(\mathbf{D}^{-1}+\overline{\mathbf{Y}}^H\mathbf{A}^{-1}\overline{\mathbf{Y}}\right)^{-1}\mathbf{D}^{-1}, \tag{47}$$

$$\begin{aligned}&\overline{\mathbf{Y}}^H(\mathbf{Y}\mathbf{Y}^H)^{-1}\mathbf{S}\left[\mathbf{S}^H(\mathbf{Y}\mathbf{Y}^H)^{-1}\mathbf{S}\right]^{-1}\mathbf{S}^H(\mathbf{Y}\mathbf{Y}^H)^{-1}\overline{\mathbf{Y}}\\&\quad=\mathbf{D}^{-1}\left[\mathbf{D}^{-1}+\overline{\mathbf{Y}}^H\mathbf{A}^{-1}\overline{\mathbf{Y}}-\mathbf{Z}^H\mathbf{B}^{-1}\mathbf{Z}\right]^{-1}\mathbf{Z}^H\mathbf{B}^{-1}\mathbf{Z}\left(\mathbf{D}^{-1}+\overline{\mathbf{Y}}^H\mathbf{A}^{-1}\overline{\mathbf{Y}}\right)^{-1}\mathbf{D}^{-1}\end{aligned} \tag{48}$$

$$\begin{aligned}&\mathbf{D}^{-1}-\overline{\mathbf{Y}}^H(\mathbf{Y}\mathbf{Y}^H)^{-1}\overline{\mathbf{Y}}+\overline{\mathbf{Y}}^H(\mathbf{Y}\mathbf{Y}^H)^{-1}\mathbf{S}\left[\mathbf{S}^H(\mathbf{Y}\mathbf{Y}^H)^{-1}\mathbf{S}\right]^{-1}\mathbf{S}^H(\mathbf{Y}\mathbf{Y}^H)^{-1}\overline{\mathbf{Y}}\\&\quad=\mathbf{D}^{-1}\left[\mathbf{D}^{-1}+\overline{\mathbf{Y}}^H\mathbf{A}^{-1}\overline{\mathbf{Y}}-\mathbf{Z}^H\mathbf{B}^{-1}\mathbf{Z}\right]^{-1}\mathbf{D}^{-1}\end{aligned} \tag{49}$$

where $\mathbf{Z} \triangleq \mathbf{S}^H\mathbf{A}^{-1}\overline{\mathbf{Y}}$ and $\mathbf{B} \triangleq \mathbf{S}^H\mathbf{A}^{-1}\mathbf{S}$ are introduced for notation simplicity. Using (45)–(49), inequality $l_2(\widehat{\mathbf{b}})>l_2(\mathbf{b})$ can be reduced to

$$\frac{[1,1](\mathbf{Z}^H\mathbf{B}^{-1}\mathbf{Z})^{-1}[1,1]^T}{[1,1]\left[\mathbf{D}^{-1}+\overline{\mathbf{Y}}^H\mathbf{A}^{-1}\overline{\mathbf{Y}}\right]^{-1}[1,1]^T}>\frac{[-1,1](\mathbf{Z}^H\mathbf{B}^{-1}\mathbf{Z})^{-1}[-1,1]^T}{[-1,1]\left[\mathbf{D}^{-1}+\overline{\mathbf{Y}}^H\mathbf{A}^{-1}\overline{\mathbf{Y}}\right]^{-1}[-1,1]^T} \tag{50}$$

where the positiveness of $|\mathbf{Z}^H\mathbf{B}^{-1}\mathbf{Z}|$ and $\left|\mathbf{D}^{-1}+\overline{\mathbf{Y}}^H\mathbf{A}^{-1}\overline{\mathbf{Y}}-\mathbf{Z}^H\mathbf{B}^{-1}\mathbf{Z}\right|$ can be easily verified[4]. Let $[[\alpha_{11},\alpha_{21}]^T,[\alpha_{12},\alpha_{22}]^T] \triangleq \overline{\mathbf{Y}}^H\mathbf{A}^{-1}\overline{\mathbf{Y}}$ and $[[\beta_{11},\beta_{21}]^T,[\beta_{12},\beta_{22}]^T] \triangleq \mathbf{Z}^H\mathbf{B}^{-1}\mathbf{Z}$. Using the ex-

[4]We note that $\overline{\mathbf{Y}}^H\mathbf{A}^{-1}\overline{\mathbf{Y}}-\mathbf{Z}^H\mathbf{B}^{-1}\mathbf{Z}$ is semi-positive definite since it can be written as $\overline{\mathbf{Y}}^H\mathbf{A}^{-1/2}\left[\mathbf{I}-\mathbf{A}^{-1/2}\mathbf{S}\left(\mathbf{S}^H\mathbf{A}^{-1/2}\mathbf{A}^{-1/2}\mathbf{S}\right)^{-1}\mathbf{S}^H\mathbf{A}^{-1/2}\right]\mathbf{A}^{-1/2}\overline{\mathbf{Y}}$.

plicit expression of the inverse of a $2 \times 2$ matrix, we can simplify (50) to

$$(\alpha_{12}+\alpha_{21})(\beta_{11}+\beta_{22}) \quad > \quad (\beta_{12}+\beta_{21})(\frac{1}{m}+\frac{1}{N-m}+\alpha_{11}+\alpha_{22}). \tag{51}$$

Set $\overline{\mathbf{z}}_1 \triangleq \overline{\mathbf{y}}_1 - \mathbf{v}$ and $\overline{\mathbf{z}}_2 \triangleq \overline{\mathbf{y}}_2 - \mathbf{v}$. Then $\overline{\mathbf{z}}_1 \sim N(\mathbf{0}, \frac{\mathbf{R}_n}{m})$ and $\overline{\mathbf{z}}_2 \sim N(\mathbf{0}, \frac{\mathbf{R}_n}{N-m})$ are independent of $\mathbf{A}$. Since $\mathbf{v}^H\mathbf{A}^{-1}\mathbf{S}\left(\mathbf{S}^H\mathbf{A}^{-1}\mathbf{S}\right)^{-1}\mathbf{S}^H = \mathbf{v}^H$, then

$$\begin{aligned}
\alpha_{11}+\alpha_{22} &= \zeta + \underbrace{\overline{\mathbf{z}}_1^H\mathbf{A}^{-1}\overline{\mathbf{z}}_1 + \overline{\mathbf{z}}_2^H\mathbf{A}^{-1}\overline{\mathbf{z}}_2}_{\delta_1}, && (52)\\
\alpha_{12}+\alpha_{21} &= \zeta + \underbrace{2\mathrm{Re}\left[\overline{\mathbf{z}}_1^H\mathbf{A}^{-1}\overline{\mathbf{z}}_2\right]}_{\delta_2}, && (53)\\
\beta_{11}+\beta_{22} &= \zeta + \underbrace{\overline{\mathbf{z}}_1^H\mathbf{A}^{-1}\mathbf{S}\left(\mathbf{S}^H\mathbf{A}^{-1}\mathbf{S}\right)^{-1}\mathbf{S}^H\mathbf{A}^{-1}\overline{\mathbf{z}}_1 + \overline{\mathbf{z}}_2^H\mathbf{A}^{-1}\mathbf{S}\left(\mathbf{S}^H\mathbf{A}^{-1}\mathbf{S}\right)^{-1}\mathbf{S}^H\mathbf{A}^{-1}\overline{\mathbf{z}}_2}_{\delta_3}, && (54)\\
\beta_{12}+\beta_{21} &= \zeta + \underbrace{2\mathrm{Re}\left[\overline{\mathbf{z}}_1^H\mathbf{A}^{-1}\mathbf{S}\left(\mathbf{S}^H\mathbf{A}^{-1}\mathbf{S}\right)^{-1}\mathbf{S}^H\mathbf{A}^{-1}\overline{\mathbf{z}}_2\right]}_{\delta_4}, && (55)\\
\zeta &= 2\mathbf{v}^H\mathbf{A}^{-1}\mathbf{v} + 2\mathrm{Re}\left[\mathbf{v}^H\mathbf{A}^{-1}\overline{\mathbf{z}}_1\right] + 2\mathrm{Re}\left[\mathbf{v}^H\mathbf{A}^{-1}\overline{\mathbf{z}}_2\right]. && (56)
\end{aligned}$$

As the SNR increases, $\delta_i, \quad i=1,\ldots,4$ are negligible comparing to $\zeta$, and (51) is asymptotically equivalent to $\zeta < 0$, i.e., $\mathrm{Re}\left[\mathbf{v}^H\mathbf{A}^{-1}\left(\overline{\mathbf{z}}_1+\overline{\mathbf{z}}_2\right)\right] < -\mathbf{v}^H\mathbf{A}^{-1}\mathbf{v}$. Given $\mathbf{A}$, $\mathbf{v}^H\mathbf{A}^{-1}\left(\overline{\mathbf{z}}_1+\overline{\mathbf{z}}_2\right)$ is a circular complex Gaussian variable with zero mean and variance $\left(\frac{1}{m}+\frac{1}{N-m}\right)\mathbf{v}^H\mathbf{A}^{-1}\mathbf{R}_n\mathbf{A}^{-1}\mathbf{v}$. Thus

$$\lim_{E\to\infty} P\left(l_2(\widehat{\mathbf{b}}) > l_2(\mathbf{b})\,\Big|\,\mathbf{A},\mathbf{b}\right) = Q\left(\sqrt{\frac{2m(N-m)\left(\mathbf{v}^H\mathbf{A}^{-1}\mathbf{v}\right)^2}{N\mathbf{v}^H\mathbf{A}^{-1}\mathbf{R}_n\mathbf{A}^{-1}\mathbf{v}}}\right). \tag{57}$$

Then (21) follows by [23] and the observation that $\frac{\left(\mathbf{v}^H\mathbf{A}^{-1}\mathbf{v}\right)^2}{\gamma(\mathbf{v}^H\mathbf{A}^{-1}\mathbf{R}_n\mathbf{A}^{-1}\mathbf{v})}$ is a scaled Beta distributed variable with parameters $a = N-L$ and $b = L-1$. ■

# 8. List of Acronyms

AFRL - Air Force Research Laboratory

AWGN - Additive White Gaussian Noise

BER - bit error rate

CDMA - code-division multiple access

CITE - Center for Integrated Transmission and Exploitation

DPSK - differential phase-shift-keying

DS-CDMA - direct sequence code-division multiple access

ICASSP - International Conference on Acoustics, Speech, and Signal Processing

IF - Information Directorate

GLRT - generalized likelihood ratio test

LMS - Least Mean Squares

LRT - likelihood ratio test

MAI - multiple access interference

MF - Matched- filter

ML - maximum likelihood

MMSE - Minimum Mean Square Estimation

RLS - Recursive Least Squares

SAB - Scientific Advisory Board

SMI - Sample Matrix Inversion

SNR - signal to noise ration

UMP - uniformly most powerful